A Bartlett-type correction for likelihood ratio tests with application to testing equality of Gaussian graphical models

Erika Banzato[1], Monica Chiogna[2], Vera Djordjilović[3], Davide Risso[1]

Statistical methods and models for complex data — Poster Session: 21 September 2022

Abstract

- We define a new correction for the likelihood ratio test (LRT) for a **two-sample problem** within the multivariate normal (MVN) context.
- The adjusted statistic, T_n, leads to **valid inference at different dimensionality** regimes of the dimension p.
- The proposed correction **improves the approximation accuracy** to the asymptotic chi-square distribution.
- T_n finds a natural application in the context of **decomposable Gaussian graphical models** (GGM).
- This renders tractable inference of **large-scale graphical models**. It improves the power of detecting a difference, and allows to **localize that difference**.

LRT for testing equality of distributions

Consider two p-dimensional MVN distributions,

$$N_p(\mu^{(j)}, \Sigma^{(j)}), \quad j = 1, 2$$

and the hypothesis of **equality of distributions** of two independent random samples of size n_j

$$H_0 : \mu^{(1)} = \mu^{(2)},\ \Sigma^{(1)} = \Sigma^{(2)} \quad \text{vs.} \quad H_a : H_0 \text{ is not true.} \tag{1}$$

The LRT for testing (1), derived in Wilks (1938), is

$$\Lambda_n = \frac{\prod_{j=1}^{2} \det\left(\hat{\Sigma}^{(j)}\right)^{n_j/2}}{\det\left(\hat{\Sigma}\right)^{n/2}},$$

where $n = n_1 + n_2$, $\hat{\Sigma}$ and $\hat{\Sigma}^{(j)}$ are the maximum likelihood estimates of the covariance matrices under H_0 and H_a.

Bartlett correction

Bartlett (1937) proposed a correction of $W_n = -2\log\Lambda_n$, that makes its mean exactly equal to the mean of the asymptotic χ^2 distribution, f. The corrected statistic takes the form

$$W_n^B = \frac{f}{E_{H_0}(W_n)} W_n, \tag{2}$$

where $E_{H_0}(W_n)$ is the expected value of W_n under H_0.

p is fixed and n is allowed to grow

Muirhead (1982) derived an expansion of the correction factor in (2),

$$\rho = 1 - \frac{2p^2 + 9p + 11}{6(p+3)n}\left(\sum_{j=1}^{2}\frac{n}{n_j} - 1\right).$$

The corrected statistic $W_n^\rho = -2\rho\log\Lambda_n$, has a χ^2 limit.

p grows at the same rate of n

Jiang and Qi (2015) established the following result based on the central limit theorem (CLT)

$$\frac{\log\Lambda_n - \mu_n}{n\sigma_n} \xrightarrow{d} N(0,1),$$

where μ_n and $\sigma_n > 0$ are the asymptotic mean and standard deviation of $\log\Lambda_n$, respectively.

Phase transition boundary

The *phase transition boundary*, d, (He et. al, 2021) characterizes the **approximation accuracy** by establishing the necessary and sufficient condition for the χ^2 approximation to hold. The condition is $p/n^d \to 0$, with

- $d = 1/2$ for W_n
- $d = 2/3$ for W_n^ρ.

Our proposal

Under the assumption that p changes with the sample size n, we propose the following adjusted statistic

$$T_n = \delta_n W_n, \quad \delta_n = \frac{f}{\mu_{w_n}}, \tag{3}$$

with $\mu_{w_n} = -2\mu_n$. The term μ_n was defined by Jiang and Qi (2015) and takes the form

$$\mu_n = [-4p - \sum_{j=1}^{2}\frac{p}{n_j} + nr_n^2(2p - 2n + 3) + - \sum_{j=1}^{2} n_j r_{n_j'}^2(2p - 2n_j + 3)]/4,$$

where $n_j' = n_j - 1$ and $r_x = (-\log(1 - p/x))^{1/2}$, for $x > p$, and $n = n_1 + n_2$.

Theorem

Let $\mathbf{p} = (p_n)_{n\in N}$ be a sequence of integers $1 \le p_n < n_j - 1$. Under H_0, for T_n defined as in (3), $\min_{j=1,2} n_j \to \infty$ and $p/n \to 0$, we have that

$$\sup_{-\infty<x<\infty} |P(T_n < x) - P(\chi_{f_n}^2 < x)| \to 0$$

and the *phase transition boundary* of T_n is $d = 1$.

Simulation study

Empirical distribution

Figure 1 shows the empirical distributions of the statistics W_n, W_n^ρ, W_n^{clt}, and T_n, along with their theoretical approximations.

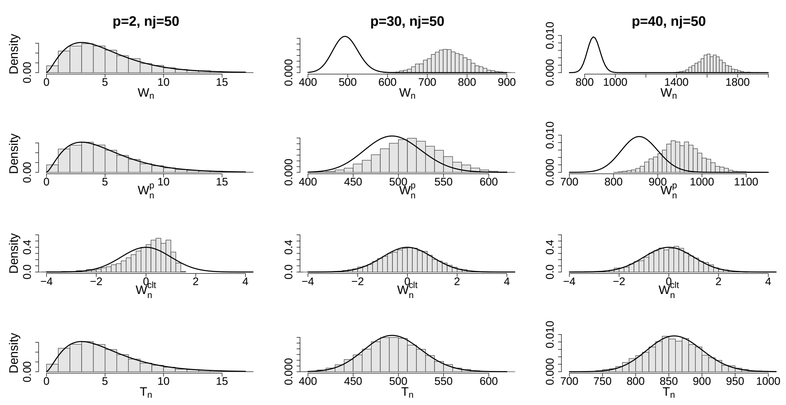

Figure 1. Simulation results with $n_1 = n_2 = 50$ and $p = 2, 30, 40$. From the top to the bottom row: empirical distribution of W_n, W_n^ρ, W_n^{clt}, and T_n. The solid line in the $1st$, $2nd$, and $4th$ rows shows the nominal χ^2 distribution, with $f = 5, 495, 860$ (from left to right). The solid line in the $3rd$ row shows the standard normal distribution.

Phase transition boundary

Figure 2 examines the *phase transition boundary* of the statistics W_n, W_n^ρ, and T_n, under the null hypothesis.

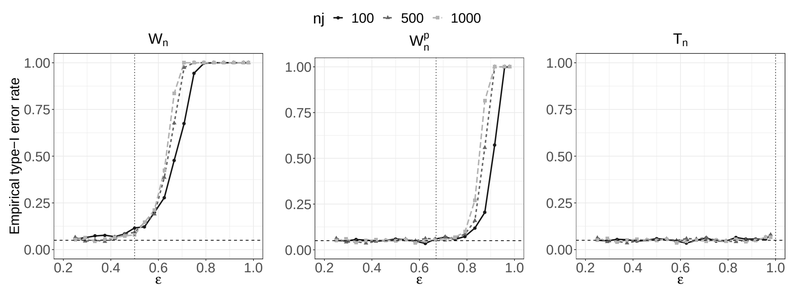

Figure 2. Empirical type-I error rate over 1000 simulations for W_n, W_n^ρ and T_n, with $p = \lfloor n_1^\varepsilon \rfloor$, $\varepsilon \in \{6/24, \ldots, 23/24, 23.5/24\}$, $n_1 = n_2$, $n = \sum_{j=1}^{2} n_j$ and $n_j \in \{100, 500, 1000\}$. The vertical dotted lines represent the phase transition boundaries for the three statistics: $1/2$, $2/3$ and 1, respectively. The horizontal dashed line represents the nominal significance level, 0.05.

Application to graphical models

Let $G = (V, E)$ denote a decomposable undirected graph, with V and E a finite set of nodes and edges. Let

- C_i, $i = 1, \ldots, k$, be a sequence of cliques satisfying the running intersection property;
- $S_i = C_i \cap C_{i-1}$, $S_1 = \varnothing$, be the set of separators;
- $R_i = C_i \setminus C_{i-1}$, $R_1 = C_1$, be the set of residuals.

According to G, the probability distribution of the random vector X_V factorizes as

$$f(X_V) = f(X_{C_1}) f(X_{R_2}|X_{S_2}) \ldots f(X_{R_k}|X_{S_k}).$$

The global hypothesis in (1) decomposes as

$$H = \bigcap_{i=1}^{k} H_i, \quad H_i : X_{R_i}^{(1)}|X_{S_i}^{(1)} \stackrel{d}{=} X_{R_i}^{(2)}|X_{S_i}^{(2)}. \tag{4}$$

The factorization of the LRT is

$$W_n = \sum_{i=1}^{k}\left[W_n^{C_i} - W_n^{S_i}\right] = W_n^{C_1} + \sum_{i=2}^{k} W_n^{C_i|S_i}.$$

The corrected statistics for the tests relating to (4) are

$$T_n^{C_i|S_i} = \delta_n^{C_i|S_i} W_n^{C_i|S_i}, \quad \delta_n^{C_i|S_i} = \frac{f_{C_i|S_i}}{\mu_n^{C_i|S_i}}, \quad \mu_n^{C|S} = \mu_n^C - \mu_n^S.$$

Simulation study

Table (1) shows the results of the tests of equality of distributions of independent samples from two conditions with respect to the graph in Figure 3.

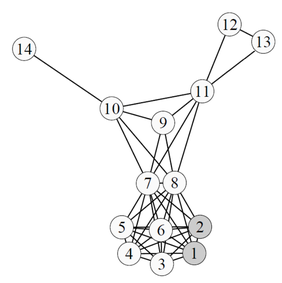

Figure 3. Nodes 1 and 2 (gray) are affected by a change in the second condition.

	W_n				T_n			
n_j	10	50	100	250	10	50	100	250
C_1	0.985	0.730	0.970	1.000	0.066	0.535	0.946	1.000
$C_2\|S_2$	0.445	0.082	0.065	0.056	0.048	0.051	0.050	0.049
$C_3\|S_3$	0.167	0.061	0.056	0.051	0.049	0.044	0.048	0.049
$C_4\|S_4$	0.109	0.060	0.051	0.057	0.047	0.052	0.048	0.055

Table 1. Empirical power and Type I error computed for each term of the decomposition. Number of rejected tests out of 10 thousand simulations, for different sample sizes, with significance level $\alpha = 0.05$.

Conclusions

- T_n shows good approximation to the χ^2 regardless of the dimension of the testing problem.
- The χ^2 approximation of T_n holds for $p/n^d \to 0$, with $d = 1$, improving the approximation accuracy of W_n.
- In GGM, T_n is able to identify the altered clique, while controlling the Type I error of the remaining local tests.

References

Bartlett, M.S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society A*, **160**, 268 – 282.

He, Y., Meng, B., Zeng, Z., Xu, G. (2021). On the phase transition of wilks' phenomenon. *Biometrika*, **108**, 741 – 748.

Jiang, T., Qi, Y. (2015). Likelihood ratio tests for high-dimensional normal distributions. *Scandinavian Journal of Statistics*, **42**, 988 – 1009.

Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Analysis*. Wiley & Sons.

Wilks, S.S. (1938). The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Annals of Mathematical Statistics*, **9**, 60 – 62.

Contact information

Erika Banzato, Ph.D. student
Department of Statistical Sciences, University of Padua, Italy
erika.banzato@phd.unipd.it

[1] Department of Statistical Sciences, University of Padua, Padua, Italy
[2] Department of Statistical Sciences, University of Bologna, Bologna, Italy
[3] Department of Economics, University Ca' Foscari of Venice, Venice, Italy